IN THE CLAIMS:

Please cancel claims 5, 8, 18, 21, 31 and 34. Please amend claims 1, 3, 6, 9, 14, 16, 19, 22, 27, 29, 32 and 35 as follows:

1. (currently amended) A method for pre-identifying implicitly defined communities including groups of pages of common interest from a collection of hyper-linked pages, wherein the communities have not been previously identified, comprising the steps of:

identifying a collection of hyperlinked pages from a plurality of sites, wherein each of the sites includes one or more hyper-linked pages;

identifying hyper-links between any two pages on a same site, wherein the same site is included within the plurality of sites;

removing the identified hyper-links between the two pages on a same site;

identifying a plurality of (i,j)-cores within the identified collection, the (i,j)-cores including a first set of hyperlinked pages and a second set of hyper-linked pages, wherein each page in the first set of hyperlinked pages points to every page in the second set of hyperlinked pages, and where i and j are the numbers of hyper-linked pages in the first set and hyper-linked pages in the second set, respectively, that appear in each of the identified (i,j)-cores ; and

expanding each of the identified (i,j)-cores into a full community, the full community being a subset of the pages regarding a particular topic.

~~identifying a plurality of community cores from the collection of hyper-linked pages, wherein the collection includes a plurality of sites with each of the sites having one or more hyper-linked pages, wherein each of the identified community cores includes first and second sets of pages, wherein each page in the first set points to every page in the second set;~~

~~removing the hyper-links between any two pages on a same site; and~~

~~expanding each identified core into a full community, the full community being a subset of the pages regarding a particular topic.~~

2. (canceled)

3. (currently amended) The method as recited in claim 1 2 further comprising the step of discarding the pages of predetermined sites.

4. (original) The method as recited in claim 1 further comprising the steps of:

finding highly similar pages that have different names;

replacing the highly similar pages with a single representative page; and

redirecting any hyper-links that pointed to one of the highly similar pages so that the redirected hyper-links now point to the representative page.

5. (canceled).

6. (currently amended) The method as recited in claim 1 6, wherein the step of discarding includes the steps of:

determining candidate fan pages, the candidate fan pages being those pointing to at least a predetermined number of different sites;

determining candidate center pages, the candidate center pages being those pointed to by one or more candidate fan pages; and

discarding all pages in the collection except the candidate fan pages and candidate center pages.

7. (original) The method as recited in claim 6, wherein the determination of candidate fan pages is based on page content and the hyper-links pointing therefrom.

8. (canceled).

9. (currently amended) The method as recited in claim 1 8, wherein the step of finding a plurality of (i, j)-cores includes the steps of:

(a) discarding all candidate center pages that have fewer than i hyper-links pointing thereto;

(b) determining all candidate center pages that have i hyper-links pointing thereto and determining whether the i hyper-links represent a valid community core; and

(c) if the i hyper-links represent a valid community core, then outputting the valid core, otherwise, discarding the determined candidate center pages.

10. (original) The method as recited in claim 9 further comprising the steps of:

(d) discarding all candidate fan pages that have fewer than j hyper-links pointing therefrom;

(e) determining all candidate fan pages that have j hyper-links pointing therefrom and determining whether the j hyper-links represent a valid community core; and

(f) if the j hyper-links represent a valid community core, then outputting the valid core, otherwise, discarding the determined candidate fan pages.

11. (original) The method as recited in claim 10 further comprising the step of repeating steps (a)-(f) until every candidate fan page has more than j hyper-links pointing therefrom and every candidate center page has more than i hyper-links pointing thereto.

12. (original) The method as recited in claim 10 further comprising the step of repeating steps (a)-(f) until a predetermined ending condition is satisfied.

13. (original) The method as recited in claim 10 further comprising the steps of:

determining all (2,j) cores by examining all pairs of candidate fan pages;

for i = 3 to n, where n is a predetermined value:

(i) finding all (i,j)-cores by examining the (i-1,j)-cores; and

(ii) for each (i-1, j)-core, determining whether any of the candidate fan

pages may be added to the (i-1, j)-core to yield a (i,j)-core; and

removing all (i,j)-cores that appear as subsets of (i',j) cores, where i' > i.

14. (currently amended) A computer program product for use with a computer

system for pre-identifying implicitly defined communities including groups of pages

of common interest from a collection of hyper-linked pages, wherein the

communities have not been previously identified, the computer program product

comprising:

a computer-readable medium;

~~means, provided on the computer-readable medium, for directing the~~

~~system to identify a plurality of community cores from the collection of hyper-~~

~~linked pages, wherein the collection includes a plurality of sites with each of~~

~~the sites having one or more hyper-linked pages, wherein each of the~~

~~identified community cores includes first and second sets of pages, wherein~~

~~each page in the first set points to every page in the second set;~~

~~means for directing the system to remove the hyper-links between any~~

~~two pages on a same site; and~~

~~means, provided on the computer-readable medium, for directing the~~

~~system to expand each identified core into a full community, the full~~

~~community being a subset of the pages regarding a particular topic.~~

means, provided on the computer-readable medium, for identifying a
collection of hyperlinked pages from a plurality of sites, wherein each of the
sites includes one or more hyper-linked pages;

means, provided on the computer-readable medium, for identifying
hyper-links between any two pages on a same site, wherein the same site is
included within the plurality of sites;

means, provided on the computer-readable medium, for removing the
identified hyper-links between the two pages on a same site;

means, provided on the computer-readable medium, for identifying a
plurality of (i,j)-cores within the identified collection, the (i,j)-cores including a
first set of hyperlinked pages and a second set of hyper-linked pages,
wherein each page in the first set of hyperlinked pages points to every page
in the second set of hyperlinked pages, and where i and j are the numbers of
hyper-linked pages in the first set and hyper-linked pages in the second set,
respectively, that appear in each of the identified (i,j)-cores ; and

means, provided on the computer-readable medium, for expanding
each of the identified (i,j)-cores into a full community, the full community
being a subset of the pages regarding a particular topic.

15. (canceled)

16. (currently amended) The computer program product as recited in claim
14~~15~~ further comprising means, provided on the computer-readable medium, for
directing the system to discard the pages of predetermined sites.

17. (original) The computer program product as recited in claim 14 further comprising:

means, provided on the computer-readable medium, for directing the system to find highly similar pages that have different names;

means, provided on the computer-readable medium, for directing the system to replace the highly similar pages with a single representative page; and

means, provided on the computer-readable medium, for directing the system to redirect any hyper-links that pointed to one of the highly similar pages so that the redirected hyper-links now point to the representative page.

18. (canceled).

19. (currently amended) The computer program product as recited in claim 14~~18~~, wherein the means for directing to discard includes:

means, provided on the computer-readable medium, for directing the system to determine candidate fan pages, the candidate fan pages being those pointing to at least a predetermined number of different sites;

means, provided on the computer-readable medium, for directing the system to determine candidate center pages, the candidate center pages being those pointed to by one or more candidate fan pages; and

means, provided on the computer-readable medium, for directing the system to discard all pages in the collection except the candidate fan pages and candidate center pages.

20. (original) The computer program product as recited in claim 19, wherein the determination of candidate fan pages is based on page content and the hyper-links pointing therefrom.

21. (canceled).

22. (currently amended) The computer program product as recited in claim 14~~21~~, wherein the means for directing to find a plurality of (i, j)-cores includes:

(a) means, provided on the computer-readable medium, for directing the system to discard all candidate center pages that have fewer than i hyper-links pointing thereto;

(b) means, provided on the computer-readable medium, for directing the system to determine all candidate center pages that have i hyper-links pointing thereto and determining whether the i hyper-links represent a valid community core; and

(c) means, provided on the computer-readable medium, for directing the system to output the valid core if the i hyper-links represent a valid community core, otherwise, to discard the determined candidate center pages.

23. (original) The computer program product as recited in claim 22 further comprising:

(d) means, provided on the computer-readable medium, for directing the system to discard all candidate fan pages that have fewer than j hyper-links pointing therefrom;

(e) means, provided on the computer-readable medium, for directing the system to determine all candidate fan pages that have j hyper-links pointing therefrom and determining whether the j hyper-links represent a valid community core; and

(f) means, provided on the computer-readable medium, for directing the system to output the valid core if the j hyper-links represent a valid community core, otherwise, discard the determined candidate fan pages.

24. (original) The computer program product as recited in claim 23, wherein the operation of means (a)-(f) is repeated until every candidate fan page has more than j hyper-links pointing therefrom and every candidate center page has more than i hyper-links pointing thereto.

25. (original) The computer program product as recited in claim 23, wherein the operation of means (a)-(f) is repeated until a predetermined ending condition is satisfied.

26. (original) The computer program product as recited in claim 23 further comprising:

means, provided on the computer-readable medium, for directing the system to determine all (2,j) cores by examining all pairs of candidate fan pages;

for i = 3 to n, where n is a predetermined value:

(i) means, provided on the computer-readable medium, for directing the system to find all (i,j)-cores by examining the (i-1,j)-cores; and

(ii) for each (i-1, j)-core, means, provided on the computer-readable medium, for directing the system to determine whether any of the candidate fan pages may be added to the (i-1, j)-core to yield a (i,j)-core; and

means, provided on the computer-readable medium, for directing the system to remove all (i,j)-cores that appear as subsets of (i',j) cores, where i' > i.

27. (currently amended) A system for pre-identifying implicitly defined communities including groups of pages of common interest from a collection of hyper-linked pages, wherein the communities have not been previously identified, comprising:

means for identifying a plurality of community cores from the collection of hyper-linked pages, wherein the collection includes a plurality of sites with each of the site having one or more hyper-linked pages, wherein each of the identified community cores includes first and second sets of pages, wherein each page in the first set points to every page in the second set;

means for removing the hyper-links between any two pages on the same site; and

means for expanding each identified core into a full community, the full

community being a subset of the pages regarding a particular topic.

means for identifying a collection of hyperlinked pages from a plurality of sites, wherein each of the sites includes one or more hyper-linked pages;

means for identifying hyper-links between any two pages on a same site, wherein the same site is included within the plurality of sites;

means for removing the identified hyper-links between the two pages on a same site;

means for identifying a plurality of (i,j)-cores within the identified collection, the (i,j)-cores including a first set of hyperlinked pages and a second set of hyper-linked pages, wherein each page in the first set of hyperlinked pages points to every page in the second set of hyperlinked pages, and where i and j are the numbers of hyper-linked pages in the first set and hyper-linked pages in the second set, respectively, that appear in each of the identified (i,j)-cores ; and

means for expanding each of the identified (i,j)-cores into a full community, the full community being a subset of the pages regarding a particular topic.

28. (canceled)

29. (currently amended) The system as recited in claim 27 28 further comprising means for discarding the pages of predetermined sites.

30. (original) The system as recited in claim 27 further comprising:

means for finding highly similar pages that have different names;

means for replacing the highly similar pages with a single representative page; and

means for redirecting any hyper-links that pointed to one of the highly similar pages so that the redirected hyper-links now point to the representative page.

31. (canceled).

32. (currently amended) The system as recited in claim 27 31, wherein the means for discarding includes:

means for determining candidate fan pages, the candidate fan pages being those pointing to at least a predetermined number of different sites;

means for determining candidate center pages, the candidate center pages being those pointed to by one or more candidate fan pages; and

means for discarding all pages in the collection except the candidate fan pages and candidate center pages.

33. (original) The system as recited in claim 32, wherein the determination of candidate fan pages is based on page content and the hyper-links pointing therefrom.

34. (canceled).

35. (currently amended) The system as recited in claim ~~27~~ 34, wherein the means for finding a plurality of (i, j)-cores includes:

(a) means for discarding all candidate center pages that have fewer than i hyper-links pointing thereto;

(b) means for determining all candidate center pages that have i hyper-links pointing thereto and determining whether the i hyper-links represent a valid community core; and

(c) means for outputting the valid core if the i hyper-links represent a valid community core, otherwise, discarding the determined candidate center pages.

36. (original) The system as recited in claim 35 further comprising:

(d) means for discarding all candidate fan pages that have fewer than j hyper-links pointing therefrom;

(e) means for determining all candidate fan pages that have j hyper-links pointing therefrom and determining whether the j hyper-links represent a valid community core; and

(f) means for outputting the valid core if the j hyper-links represent a valid community core, otherwise, discarding the determined candidate fan pages.

37. (original) The system as recited in claim 36, wherein the operation of means (a)-(f) is repeated until every candidate fan page has more than j hyper-links

pointing therefrom and every candidate center page has more than i hyper-links

pointing thereto.

38. (original) The system as recited in claim 36, wherein the operation of

means (a)-(f) is repeated until a predetermined ending condition is satisfied.

39. (original) The system as recited in claim 36 further comprising:

means for determining all (2,j) cores by examining all pairs of candidate fan pages;

for i = 3 to n, where n is a predetermined value:

(i) means for finding all (i,j)-cores by examining the (i-1,j)-cores; and

(ii) for each (i-1, j)-core, means for determining whether any of the candidate fan pages may be added to the (i-1, j)-core to yield a (i,j)-core; and

means for removing all (i,j)-cores that appear as subsets of (i',j) cores, where i' > i.